

粒子群算法稳定性的参数选择策略分析

张 玮^{1,2}, 王华奎²

(1. 太原理工大学化学化工学院, 太原 030024; 2. 太原理工大学信息工程学院, 太原 030024)



摘 要: 在对标准粒子群算法位置期望及位置方差进行稳定性分析的基础上, 研究加速因子的取值对位置期望及位置方差的影响。从提高系统稳定性的角度考虑, 建议个体认知加速因子为 1.85, 社会认知加速因子为 2, 通过五个经典测试函数与 Kennedy 提出的设定两个加速因子均为 2 的建议做比较。大量实验统计结果表明对于同一种算法本文建议的参数设置在寻优速度和方差上均优于 Kennedy 提出的参数设置, 更有利于工程应用。

关键词: 粒子群算法; 标准差; 稳定性; 参数选择

中图分类号: TP301.6

文献标识码: A

文章编号: 1004-731X (2009) 14-4339-06

Analysis on Parameter Selection
for Stability of Particle Swarm Optimization AlgorithmZHANG Wei^{1,2}, WANG Hua-kui²

(1. College of Chemistry and Chemical Engineering, Taiyuan University of Technology, Taiyuan 030024, China;

2. College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: The effect of acceleration factors on expectation and standard deviation values of the position was studied based on the stability analysis of Particle Swarm Optimization Algorithm. From the viewpoint of improving system stability, setting the cognitive acceleration factor as 1.85 and the social acceleration factor as 2 is a better choice than that Kennedy proposed both as 2. Moreover, five benchmark functions were used to test the influence of these parameters on system stability in the same strategy with numerous experiments. The statistical results from the new strategy of parameters selection also show higher efficiency than that Kennedy suggested.

Key words: particle swarm optimization algorithm; standard deviation; stability; parameter selection

引 言

粒子群优化算法(Particle Swarm Optimization, PSO)是由 Kennedy 和 Eberhart 于 1995 年提出基于群体智能的随机优化策略^[1]。PSO 概念简单、参数较少、收敛速度快, 目前已有较深入的理论和较广泛的应用。许多学者对标准粒子群算法的收敛性进行了研究: 1999 年 Ozcan 和 Mohan 对简单粒子群系统首次进行了理论分析^[2]; 2001 年 van den Bergh^[3]引用 Solis 和 Wets 关于随机性算法的收敛准则, 证明了标准 PSO 算法收敛于全局最优解与局部最优解线性关系之间的某一稳定点, 并指出保证收敛的参数取值范围; 2002 年 Clerc 和 Kennedy^[4]对多维复杂空间中的粒子群算法的稳定性和收敛性进行了理论分析; 2003 年 Trelea^[5]在对标准粒子群算法收敛性分析的基础上提出了收敛性和稳定性较好的一组参数选择; 2006 年 M. Jiang^[6]对标准粒子群算法进行了随机收敛分析, 提出了算法收敛的条件及参数选择范围。以上研究的重点都集中于研究 PSO 算法的收敛域及对应的参数选择。

笔者认为虽然算法在收敛域内必定要收敛, 但收敛速度

太低或稳定性太差均不利于工程应用。因此, 研究标准粒子群算法在收敛域内的收敛效率, 提高随机算法的稳定性是十分必要的。本文第 2 节分析了粒子运动位置期望及方差的稳定性, 稳态结果表明二者的大小与加速因子有关; 第 3 节进一步研究加速因子的取值对位置期望及方差大小的影响, 并从减小方差提高系统稳定性的角度建议个体认知加速因子为 1.85, 社会认知加速因子为 2; 第 4 节对五个函数采用同一策略的不同参数取值进行寻优比较得出本文所提参数的有效性。

1 标准粒子群优化算法

标准粒子群算法的迭代方程^{[3][7]}如下:

$$v_{id}^{t+1} = \omega \times v_{id}^t + c_1 \times r1_{id}^t \times (p_{id}^t - x_{id}^t) + c_2 \times r2_{id}^t \times (p_{gd}^t - x_{id}^t) \quad (1)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \quad (2)$$

其中 $d=1,2,\dots,D$ 是目标搜索空间的维数, $i=1,2,\dots,N$ 是粒子个数, v_{id}^t 是粒子 i 在第 t 次迭代中第 d 维的速度, x_{id}^t 是粒子 i 在第 t 次迭代中第 d 维的位置, p_{id}^t 是粒子 i 在第 t 次迭代中第 d 维的个体最优位置, p_{gd}^t 是粒子 i 在第 t 次迭代中第 d 维的全局最优位置, $r1$ 和 $r2$ 为均匀分布于 $[0,1]$ 之间的随机数, c_1 和 c_2 为加速因子, 用来调节每次迭代的步长, ω 为惯性因子, ω 较大则全局搜索能力强, 局部搜索能力弱, 反之, 则局部搜索能力增强, 而全局搜索能力减弱。

收稿日期: 2008-02-18 **修回日期:** 2008-04-23

作者简介: 张玮(1973-), 女, 山西沁水人, 博士生, 讲师, 研究方向为智能优化算法的理论及应用; 王华奎(1946-), 男, 山西襄垣人, 教授, 博导, 研究方向为智能信号处理, 超宽带通信。

2 PSO 稳定性分析

作为一种随机搜索优化方法, PSO 的稳定性是一项重要的指标。在粒子系统进化的过程中, p_i 和 p_g 是经常更新的。为了分析方便, 可以考虑在一个很短的时间内 p_i 和 p_g 保持为常量。这样, 只需研究粒子 i 的运动就可以了。由于粒子 i 是任意选择的, 因此其结果可以适用于其他所有的粒子。同时, 由迭代公式 (1) 和 (2) 可以看出, 每一维的更新是相互独立的, 这样算法就可以降为一维情况进行研究^[6]。忽略粒子 i 和维数符号 d , 离散状态更新公式可以重新写为:

$$\begin{cases} v_{t+1} = \omega v_t + c_1 r_{1,t} (p_i - x_t) + c_2 r_{2,t} (p_g - x_t) \\ x_{t+1} = x_t + v_{t+1} \end{cases} \quad (3)$$

$$x_{t+1} = x_t + v_{t+1} \quad (4)$$

2.1 位置期望的稳定性分析

把速度公式 (3) 带入位置公式 (4), 得到位置的离散迭代公式, 整理得:

$$x_{t+1} + (c_1 r_{1,t} + c_2 r_{2,t} - 1 - \omega)x_t + \omega x_{t-1} - c_1 r_{1,t} p_i - c_2 r_{2,t} p_g = 0 \quad (5)$$

令初始位置为 x_0 , 初始速度为 v_0 , 由公式 (5) 可知第 t 代位置 x_t 是一个随机变量, 从而 $\{x_t\}$ 即为随机序列, 令此序列的数学期望为 $\{Ex_t\}$, 由于 r_1 和 r_2 均匀分布于 $[0, 1]$ 之间, 则 $E(r_1) = E(r_2) = 1/2$, 将其代入公式 (5), 位置期望的迭代公式可以写为:

$$Ex_{t+2} + \left(\frac{c_1 + c_2}{2} - 1 - \omega\right)Ex_{t+1} + \omega Ex_t - \frac{c_1 p_i + c_2 p_g}{2} = 0 \quad (6)$$

根据实数位移定理对差分方程 (6) 进行 Z 变换^[7], 整理得:

$$Ex(z) = \frac{z^2 x_0 + z x_1 + \left(\frac{c_1 + c_2}{2} - \omega - 1\right) z x_0 + \frac{c_1 p_i + c_2 p_g}{2} \frac{z}{z-1}}{z^2 + \left(\frac{c_1 + c_2}{2} - \omega - 1\right) z + \omega} \quad (7)$$

其特征方程为:

$$z^2 + \left(\frac{c_1 + c_2}{2} - \omega - 1\right)z + \omega = 0 \quad (8)$$

离散系统稳定的充分必要条件是: 系统特征方程的所有根均位于 z 平面上以原点为圆心的单位圆之内, 也就是特征方程的模均小于 1。为了分析位置期望的稳定性, 对其特征方程 (8) 做双线性变换。

令 $z = \frac{u+1}{u-1}$ 代入 (8), 整理得

$$(c_1 + c_2)u^2 + 4(1 - \omega)u + 4 + 4\omega - (c_1 + c_2) = 0 \quad (9)$$

列 Routh 表

$$\begin{array}{ccc} u^2 & c_1 + c_2 & 4 + 4\omega - (c_1 + c_2) \\ u^1 & 4(1 - \omega) & \\ u^0 & 4 + 4\omega - (c_1 + c_2) & \end{array}$$

根据离散系统稳定的充要条件, 令 Routh 表的第一列元素均大于零

$$\begin{cases} c_1 + c_2 > 0 \\ 4(1 - \omega) > 0 \\ 4 + 4\omega - (c_1 + c_2) > 0 \end{cases} \Rightarrow \begin{cases} -1 < \omega < 1 \\ 0 < c_1 + c_2 < 4(1 + \omega) \end{cases} \quad (10)$$

条件(10)就是序列 $\{Ex_t\}$ 稳定的充要条件。

若位置的期望是稳定的, 则可用 z 变换的终值定理求出位置的稳态值

$$Ex(\infty) = \lim_{z \rightarrow 1} (z-1)Ex(z) = \frac{c_1 p_i + c_2 p_g}{c_1 + c_2} \quad (11)$$

由此可见, 终值是在 p_i 与 p_g 之间的某一值, 并且 c_2 越大, 终值就越接近 p_g 。France van den Bergh 以及其他学者已用不同的方法证明了此结果。

2.2 位置方差的稳定性分析

重写位置迭代公式 (5) 为

$$x_{t+1} = (1 + \omega - c_1 r_{1,t} - c_2 r_{2,t})x_t - \omega x_{t-1} + c_1 r_{1,t} p_i + c_2 r_{2,t} p_g \quad (12)$$

对(12)式两边取方差, 并整理得

$$\begin{aligned} Dx_{t+2} - \left(1 + \omega^2 - \frac{c_1^2 + c_2^2}{3}\right)Dx_{t+1} + \omega^2 Dx_t - \\ \frac{1}{12} \left[c_1^2 p_i^2 + c_2^2 p_g^2 - \frac{c_1^2 + c_2^2}{c_1 + c_2} (c_1 p_i + c_2 p_g)^2 \right] = 0 \end{aligned} \quad (13)$$

对 (13) 式进行 z 变换

$$Dx(z) = \frac{-\frac{1}{12} \left[c_1^2 p_i^2 + c_2^2 p_g^2 - \frac{c_1^2 + c_2^2}{c_1 + c_2} (c_1 p_i + c_2 p_g)^2 \right] \frac{z}{z-1}}{z^2 - \left(1 + \omega^2 - \frac{c_1^2 + c_2^2}{3}\right)z + \omega^2} \quad (14)$$

为了分析稳定性, 对 (14) 式的特征方程进行双线性变换并整理得:

$$\left(\frac{c_1^2 + c_2^2}{3} - \omega^2\right)u^2 + 2u + 2 + 2\omega^2 - \frac{c_1^2 + c_2^2}{3} = 0 \quad (15)$$

列 Routh 表

$$\begin{array}{ccc} u^2 & \frac{c_1^2 + c_2^2}{3} - \omega^2 & 2 + 2\omega^2 - \frac{c_1^2 + c_2^2}{3} \\ u^1 & 2 & \\ u^0 & 2 + 2\omega^2 - \frac{c_1^2 + c_2^2}{3} & \end{array}$$

系统稳定的充要条件是 Routh 表的第一列必须全部大于零, 即

$$\begin{cases} \frac{c_1^2 + c_2^2}{3} - \omega^2 > 0 \\ 2 + 2\omega^2 - \frac{c_1^2 + c_2^2}{3} > 0 \end{cases} \Rightarrow \omega^2 < \frac{c_1^2 + c_2^2}{3} < 2 + 2\omega^2 \quad (16)$$

此式的得出应该在位置期望稳定的前提下, 因此结合位置期望稳定的条件得出位置方差稳定的最终条件是

$$\begin{cases} \omega^2 < \frac{c_1^2 + c_2^2}{3} < 2 + 2\omega^2 \\ -1 < \omega < 1 \\ 0 < c_1 + c_2 < 4(1 + \omega) \end{cases} \quad (17)$$

利用 z 变换的终值定理求得方差的稳态值为:

$$\begin{aligned} Dx(\infty) &= \lim_{z \rightarrow 1} (z-1)Dx(z) \\ &= \frac{1}{4} \left[\frac{c_1 p_i + c_2 p_g}{c_1 + c_2} - \frac{c_1^2 p_i^2 + c_2^2 p_g^2}{c_1^2 + c_2^2} \right] \end{aligned} \quad (18)$$

结论:由位置期望的稳态公式(11)和方差的稳态公式(18)可以看出其大小均与 c_1 和 c_2 有关而与惯性权值 ω 无关。但通过以上分析得出位置期望与位置方差稳定的条件与 ω 有关。惯性因子 ω 决定了搜索能力的强弱,而加速因子 c_1 和 c_2 则影响最优位置和方差的大小。从工程应用的角度来看,方差越小,寻优结果就越集中,寻优系统的稳定性就越好。因此,在系统能够收敛到最优值的前提下,寻找适当的加速因子值,使得寻优结果越集中是本文研究的重点。

3 加速因子 c_1 和 c_2 对寻优迭代的影响

令速度公式为 $v_{t+1} = V_1 + V_2 + V_3$, 其中,

$$\begin{aligned} V_1 &= \omega v_t \\ V_2 &= c_1 r_{1,t} (p_i - x_t) \\ V_3 &= c_2 r_{2,t} (p_g - x_t) \end{aligned}$$

V_1 决定了搜索能力的强弱, V_2 和 V_3 决定了寻优结果和方差的大小。为了研究方便,假设粒子飞行无记忆性,即令 $V_1 = \omega v_t = 0$,则粒子的速度由 V_2 和 V_3 决定,即粒子的速度与当前位置与局部最优位置之差、当前位置与全局最优位置之差以及 c_1 和 c_2 的取值有关。由于 V_2 和 V_3 结构相似,因此只研究一个通式中系数 c 的作用即可,此通式可写为

$$\begin{cases} v_{t+1} = cr_t(p - x_t) \\ x_{t+1} = x_t + v_{t+1} \end{cases} \quad (19)$$

其中, c 为加速因子, p 为最优值,则位置的迭代公式可写为:

$$x_{t+1} = x_t + cr_t(p - x_t) \quad (20)$$

令初始位置为 x_0 , 则

$$\begin{aligned} x_1 &= x_0(1 - cr_0) + cr_0p \\ x_2 &= x_0(1 - cr_0)(1 - cr_1) + cr_0p(1 - cr_1) + cr_1p \\ x_3 &= x_0(1 - cr_0)(1 - cr_1)(1 - cr_2) + \\ &\quad cr_0p(1 - cr_1)(1 - cr_2) + cr_1p(1 - cr_2) + cr_2p \\ x_4 &= x_0(1 - cr_0)(1 - cr_1)(1 - cr_2)(1 - cr_3) + \\ &\quad cr_0p(1 - cr_1)(1 - cr_2)(1 - cr_3) + cr_1p(1 - cr_2) \\ &\quad (1 - cr_3) + cr_2p(1 - cr_3) + cr_3p \\ &\vdots \end{aligned} \quad (21)$$

相应位置的期望为

$$\begin{aligned} Ex_1 &= x_0 \left(1 - \frac{1}{2}c\right) + \frac{1}{2}cp \\ Ex_2 &= x_0 \left(1 - \frac{2}{2}c + \frac{1}{4}c^2\right) + \left(\frac{2}{2} - \frac{1}{4}c\right)cp \\ Ex_3 &= x_0 \left(1 - \frac{3}{2}c + \frac{3}{4}c^2 - \frac{1}{8}c^3\right) + \left(\frac{3}{2} - \frac{3}{4}c + \frac{1}{8}c^2\right)cp \\ Ex_4 &= x_0 \left(1 - \frac{4}{2}c + \frac{6}{4}c^2 - \frac{4}{8}c^3 + \frac{1}{16}c^4\right) + \left(\frac{4}{2} - \frac{6}{4}c + \frac{4}{8}c^2 - \frac{1}{16}c^3\right)cp \\ &\vdots \end{aligned} \quad (22)$$

当加速因子 c 取固定值时,列表比较位置期望的变化,如表 1 所示。

由表 1 中前四次位置期望迭代的变化看出:无论 c 为何值,随着迭代次数的增加,位置期望终究都会收敛到最优值 p ; $c=2$ 时,位置的期望保持不变,始终是 p 。不难理解, Kennedy^[1]建议 $c_1=c_2=2$ 。 c 越接近于 2,初始位置 x_0 对迭代过程的影响越小,寻优的多样性降低; c 越远离 2,初始位置 x_0 对迭代过程影响越大,越能够体现寻优的多样性。根据对象的特点,选择恰当的 c 值:对于复杂度高的对象或系统陷入局部最优时,都易选择较小的 c 值;对于较简单的对象,易选择较大的 c 值。一般情况, c 选固定值就能够保证系统收敛了。

但要了解迭代过程的位置在最优值周围的聚集程度,则需要进一步研究位置的方差。

仍以通式(19)为例研究。根据(21)式的位置迭代公式,逐一求各代位置的方差。

$$\begin{aligned} Dx_1 &= D(x_0(1 - cr_0) + cr_0p) = Dx_0 - D(cx_0r_0) + D(cpr_0) \\ &= \frac{c^2}{12}(p^2 - x_0^2) \end{aligned} \quad (23)$$

$$\begin{aligned} Dx_2 &= D(x_0(1 - cr_0)(1 - cr_1) + cr_0p(1 - cr_1) + cr_1p) \\ &= D(x_0(1 - cr_0 - cr_1 + c^2r_0r_1) + p(cr_0 + cr_1 - c^2r_0r_1)) \\ &= c^2(p^2 - x_0^2)D(r_0 + r_1 - cr_0r_1) \\ &= c^2(p^2 - x_0^2)\left(\frac{1}{6} - \frac{7}{144}c^2\right) \end{aligned} \quad (24)$$

$$\text{当 } c=1.85164 \text{ 时, } \frac{1}{6} - \frac{7}{144}c^2 = 0, \text{ 即 } Dx_2 = 0$$

根据方差与期望稳定的条件公式(17)得出对应的惯性因子取值范围是

$$0.375 < \omega < 1 \quad (25)$$

$$\begin{aligned} Dx_3 &= D(x_0(1 - cr_0)(1 - cr_1)(1 - cr_2) + cr_0p(1 - cr_1)(1 - cr_2) + \\ &\quad cr_1p(1 - cr_2) + cr_2p) \\ &= c^2(p^2 - x_0^2)D(r_0 + r_1 + r_2 - cr_0r_1 - cr_0r_2 - cr_1r_2 + c^2r_0r_1r_2) \\ &= c^2(p^2 - x_0^2)\left(\frac{1}{4} - \frac{7}{48}c^2 + \frac{37}{1728}c^4\right) \end{aligned} \quad (26)$$

$$\text{当 } c=1.8454 \text{ 时, } \frac{1}{4} - \frac{7}{48}c^2 + \frac{37}{1728}c^4 \text{ 取最小值 } 0.00169,$$

$$\text{则 } \min(Dx_3) = 0.00575(p^2 - x_0^2)$$

根据方差与期望稳定的范围公式(17)得对应的惯性因子取值范围是

$$0.367 < \omega < 1 \quad (27)$$

表 1 加速因子 c 取固定值时对每次迭代位置期望的影响

	$c=2.4$	$c=2$	$c=1.6$	$c=1.2$	$c=0.6$
Ex_1	$-0.2x_0+1.2p$	p	$0.2x_0+0.8p$	$0.4x_0+0.6p$	$0.7x_0+0.3p$
Ex_2	$0.04x_0+0.96p$	p	$0.04x_0+0.96p$	$0.16x_0+0.84p$	$0.49x_0+0.51p$
Ex_3	$-0.008x_0+1.008p$	p	$0.008x_0+0.992p$	$0.064x_0+0.936p$	$0.343x_0+0.657p$
Ex_4	$0.0016x_0+0.9984p$	p	$0.0016x_0+0.9984p$	$0.0256x_0+0.9744p$	$0.2401x_0+0.7599p$

$$\begin{aligned}
 Dx_4 &= D(x_0(1-cr_0)(1-cr_1)(1-cr_2)(1-cr_3) + \\
 &\quad cr_0p(1-cr_1)(1-cr_2)(1-cr_3) + \\
 &\quad cr_1p(1-cr_2)(1-cr_3) + cr_2p(1-cr_3) + cr_3p) \\
 &= c^2(p^2 - x_0^2)D(r_0 + r_1 + r_2 + r_3 - cr_0r_1 - cr_0r_2 - cr_0r_3 - \\
 &\quad cr_1r_2 - cr_1r_3 - cr_2r_3 + c^2r_0r_1r_2 + c^2r_0r_1r_3 + c^2r_0r_2r_3 + \\
 &\quad c^2r_1r_2r_3 - c^3r_0r_1r_2r_3) \\
 &= c^2(p^2 - x_0^2)\left(\frac{1}{3} - \frac{7}{24}c^2 + \frac{37}{288}c^4 - \frac{175}{20736}c^6\right) \quad (28)
 \end{aligned}$$

$$\text{令 } Dx_i = (p^2 - x_0^2)\lambda_i(c), \quad i=1,2,3,4 \quad (29)$$

其中 $\lambda_i(c)$ 为方差系数。当 c 取不同值时, $\lambda_i(c)$ 如表 2 所示。

由表中 $\lambda_i(c)$ 的大小随迭代的变化可以看出, c 的取值越小对应的方差系数就越小, 但 c 取很小的值并不合适, 因为从表 1 可知, c 小会增加迭代的次数, 使得收敛效率降低。值得注意的是, 当 $c=1.85$ 时, $\lambda_2(c)$ 和 $\lambda_3(c)$ 都近似为零, 其对应的方差 $Dx_2 \approx Dx_3 \approx 0$, 此时寻优的随机位置越集中。

综合以上分析, 根据表 2 建议 $c_1=1.85$, 使得寻优位置与局部最优值的方差小, 减小系统在寻优过程中的振荡; 在此基础上根据表 1 可以令 $c_2=2$, 使得寻优位置期望与全局最优值最接近, 加快寻优速度。对应的惯性因子 ω 的取值范围为 $(0.375, 1)$ 。

4 实验与讨论

4.1 测试函数的定义

采用广泛应用于进化计算的五种函数^[4,8]测试 PSO 选择不同参数时的算法性能。测试函数的定义、维数、取值范围、理论极值及误差目标列于表 3。根据[9]和[10]设置粒子数为 30, $V_{max}=0.1X_{max}$ ^[12], 最大寻优次数为 5000 次。

表 2 c 取不同值时方差系数 $\lambda_i(c)$ 的变化

	$c=2.2$	$c=2$	$c=1.85$	$c=1$	$c=1.494$	$c=0.6$	$c=0.2$
$\lambda_1(c)$	0.4	0.33	0.285	0.083	0.186	0.03	0.0033
$\lambda_2(c)$	-0.33	-0.11	0.001	0.118	0.13	0.05	0.0066
$\lambda_3(c)$	0.22	0.037	0.0058	0.126	0.07	0.072	0.01
$\lambda_4(c)$	4.72	2.73	1.72	0.16	0.51	0.09	0.013

表 3 测试函数的定义

函数名	函数表达式	维数	取值范围	理论极值	误差目标
Sphere	$f(x) = \sum_{i=1}^n x_i^2$	30	$[-100, 100]^n$	0	10^{-2}
Rosenbrock	$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	$[-30, 30]^n$	0	10^2
Rastrigin	$f(x) = \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10)$	30	$[-5.12, 5.12]^n$	0	10^2
Griewank	$f(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	$[-600, 600]^n$	0	10^{-1}
Schaffer's f6	$f(x) = 0.5 + \frac{(\sin \sqrt{x_1^2 + x_2^2})^2 - 0.5}{(1 + 0.001(x_1^2 + x_2^2))^2}$	2	$[-100, 100]^n$	-1	10^{-5}

4.2 寻优策略及数值实验

4.2.1 寻优策略简介

为了验证本文所提出参数选择的有效性, 分别在线性递减惯性权策略和本文所提出的非线性递减惯性权策略中采用本文所提参数设置 $c_1=1.85$, $c_2=2$ 与 Kennedy 提出的参数设置 $c_1=c_2=2$ 作比较, 在同样的 MATLAB 7.0 环境下进行计算机仿真并与压缩因子法进行对比。

算法简介如下:

1) 线性递减惯性权策略 (Linearly Decreasing Weight, LDW)

Shi, Y.H. 建议采用线性递减惯性权策略来控制惯性权重^[8], 其变化公式为:

$$\omega(t) = \omega_{mi} + \frac{\omega_{fn} - \omega_{mi}}{t_{max}} \cdot t \quad (30)$$

式中: $\omega(t)$ 为第 t 代的惯性权重值。 t 为当前迭代次数; t_{max} 为最大迭代次数。 ω_{mi} 为初始惯性权重, ω_{fn} 为终了惯性权重。文献[9]经过大量数值实验, 推荐 $\omega_{mi}=0.9$, $\omega_{fn}=0.4$ 。实验中当迭代次数超过 1500 代时, 惯性权值取初始值。

2) 非线性递减惯性权策略 (Nonlinearly Decreasing Weight, NLDW)

虽然线性递减惯性权策略简单, 但在很多情况下收敛速度并不快, 大量的实验发现: 对于较简单的测试函数或问题采用粒子群算法求解时, 在迭代之初易选择较大的 $|\omega|$ 值, 使粒子以较快的飞行速度在整个搜索空间寻优, 以确定最优值的大致范围, 但持续时间不宜过长; 在接近寻优终了时, 易选择较小的 $|\omega|$ 值, 使粒子在最优值附近精确搜寻, 持续时间适当长一些。对于较复杂的测试函数或问题求解, 在迭代之

初选择较大的 $|\omega|$ 值,持续时间适当长一些,在接近寻优终了时,选择较小的 $|\omega|$ 值,持续时间适当短一些。因此, ω 非线性递减是一个较好的选择,递减指数函数表达式如下:

$$\omega(t) = \left(\frac{t}{N}\right)^m \cdot (\omega_{fin} - \omega_{ini}) + \omega_{ini} \quad (31)$$

其中, m 为递减指数,根据不同的寻优任务选择不同的 m 值; N 为迭代阈值,一般可取 1500。当迭代次数达到 N 时令 $\omega(t) = \omega_{ini}$,并将保持到搜索结束。根据式(25)和(27),设定 $\omega_{ini} = 0.99, \omega_{fin} = 0.38$ 。

3) 压缩因子法(Constriction Factor Method, CFM)

Clerc M.1999年^[11]提出压缩因子法,其速度更新方程为

$$v_{id}^{t+1} = \chi(v_{id}^t + \varphi_1 \times rand \times (p_{id}^t - x_{id}^t) + \varphi_2 \times rand \times (p_{gd}^t - x_{id}^t)) \quad (32)$$

其中 $\chi = \frac{2}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}}$ 为压缩因子, $\varphi = \varphi_1 + \varphi_2; \varphi > 4$ 。

典型取值 $\varphi = 4.1$,这时 $\chi = 0.729$,根据 Kennedy 和 Eberhart 的研究工作和大量的计算机仿真经验,取 $\varphi_1 = \varphi_2 = 2.05$ 。

4.2.2 实验方案设计及数据

为了研究方便,设定实验系列 1 与实验系列 2,详细说明如下:

- 1) LDW1(线性递减策略): $c_1 = c_2 = 2, \omega \in [0.4, 0.9]$
- 2) LDW2(线性递减策略): $c_1 = 1.85, c_2 = 2, \omega \in [0.4, 0.9]$
- 3) NLDW1(非线性递减策略): $c_1 = c_2 = 2, \omega \in [0.38, 0.99]$
- 4) NLDW2(非线性递减策略): $c_1 = 1.85, c_2 = 2, \omega \in [0.38, 0.99]$

5) CFM (压缩因子法)

根据经验,针对不同的测试函数选定非线性递减指数

如表 4 所示。

表 4 非线性递减指数 m 的取值

测试函数名	Sphere	Rosenbrock	Rastrigin	Griewank	Schaffer's f6
递减指数 m	0.1	0.1	0.01	0.1	3

分别对五种寻优策略进行 50 次独立仿真实验,达到误差目标要求即停止运行。仿真曲线截图,如图 1-图 5 所示。图中标记 1 为 LDW1(线性递减策略),标记 2 为 LDW2,标记 3 为 NLDW 1(非线性递减策略),标记 4 为 NLDW2,标记 5 为 CFM (压缩因子法)。

记录成功收敛率 (Suc.), 最小迭代次数(min.), 最大迭代次数 (max.) 及平均迭代次数 (Mean), 并计算标准差 (Standard Deviation, St.D), 最后根据标准差的大小排序 (Sequence)。仿真实验数据见表 5。

从图 1 至图 4 的曲线明显地看出,采用本文提出的惯性权重非线性递减策略所需的寻优迭代次数要远远小于惯性权重线性递减策略。但这并不是本文的研究重点。本文研究的重点在于,对于同样的寻优策略,采用本文提出的参数选择策略,其寻优的稳定性会有所提高,即迭代次数统计数据的方差要小。这一点在表 5 中能够看出。另外,迭代过程稳定性的提高,会带来迭代次数的降低,图 1 至图 4 能够比较直观地说明这一点,即实验系列 2 所需的迭代次数要小于实验系列 1 的迭代次数。对于图 5, Schaffer's f6 函数的优化,虽然非线性递减策略的寻优迭代数大于线性递减策略,但其迭代次数的方差要远远小于线性递减策略,即采用非线性递减策略系统的稳定性有所提高(这一点从表 5 种可以看出)。

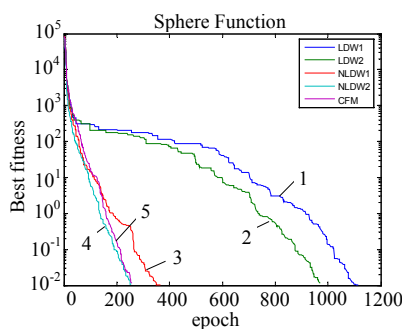


图 1 Sphere function

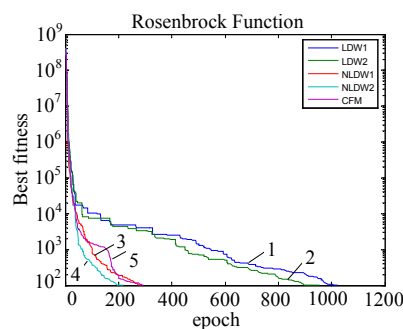


图 2 Rosenbrock function

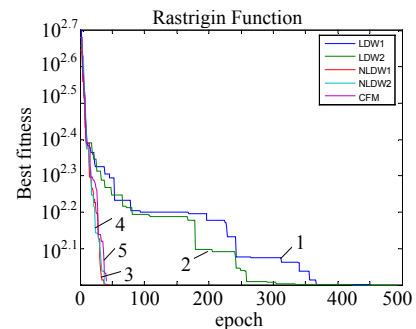


图 3 Rastrigin function

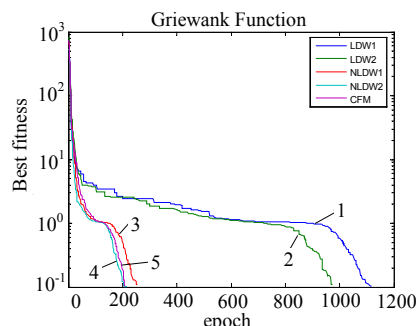


图 4 Griewank function

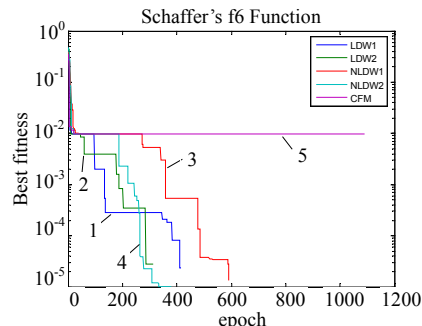


图 5 Schaffer's f6 function

表 5 仿真实验数据

Function	type	Suc.	min	max	mean	St.D	Sequence
Sphere	LDW1	1	1021	1239	1120	45.7	5
	LDW2	1	916	1062	990	33.8	3
	NLDW 1	1	268	425	331	27.2	2
	NLDW 2	1	236	347	280	26.7	1
	CFM	1	224	421	295	39.6	4
Rosenbrock	LDW1	1	754	2673	1050	279	4
	LDW2	1	709	1531	918	147	1
	NLDW 1	1	165	2054	323	289	5
	NLDW 2	1	130	1083	259	196	3
	CFM	1	131	1471	232	193	2
Rastrigin	LDW1	1	88	659	346	114	5
	LDW2	1	100	453	296	83	4
	NLDW 1	1	25	208	40	27	3
	NLDW 2	1	24	51	33	5.1	1
	CFM	1	24	123	44	16	2
Griewank	LDW1	1	951	1299	1072	57	5
	LDW2	1	855	1099	958	51	3
	NLDW 1	1	191	304	240	24	2
	NLDW 2	1	183	260	211	19	1
	CFM	0.98	177	496	260	55	4
Schaffer's f6	LDW1	0.94	140	851	295	162	4
	LDW2	0.98	147	781	277	115	3
	NLDW 1	1	304	693	409	71	2
	NLDW 2	0.98	295	593	378	54	1
	CFM	0.74	53	697	237	194	5

5 结论

从以上五种测试函数的实验数据可以得出下列结论:

1) 对于线性递减惯性权策略采用本文提出的加速因子参数设置,其迭代次数的标准差和均值均比 Kennedy 提出的参数设置有所减少。说明本文提出参数的有效性。

2) 对于非线性递减惯性权策略采用本文提出的加速因子参数设置,其迭代的标准差和均值均小于 Kennedy 提出的参数设置。更进一步说明本文提出的参数选择在同一种寻优策略下能有效提高系统稳定性。

3) 对于压缩因子法,在同样的实验条件下,其迭代均值都很小,即收敛速度很快,但其标准差的大小相对其它算法并不稳定:对于 Sphere、Griewank 和 Schaffer's f6 函数,其标准差较大;对于 Rosenbrock 和 Rastrigin 函数,其标准差较小。因此研究提高系统稳定性的策略是本文工作的意义所在。

4) 从表 5 的标准差排序中还可以看出,标准差最小的参数选择就是本文所提出的参数选择策略。

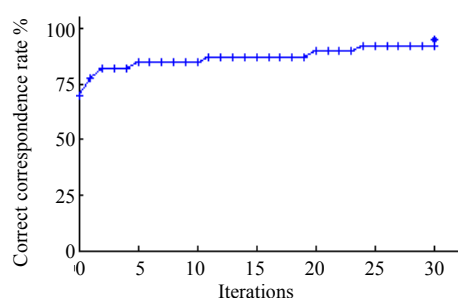
5) 采用本文提出的 NLDW 策略迭代次数的期望比 LDW 有明显降低,对于函数 Sphere, Rastrigin 和 Griewank 其迭代均值和标准差均小于 LDW,这是由于恰当地选择了非线性指数;对于函数 Rosenbrock,其迭代均值的优势明显,但标准差均比 LDW 和 CFM 大;对于函数 Schaffer's f6,其

标准差的优势明显,但迭代均值却很大。笔者下一步的工作将进一步研究非线性递减惯性权重策略的指数参数选择对系统稳定性的影响,优化其寻优性能。

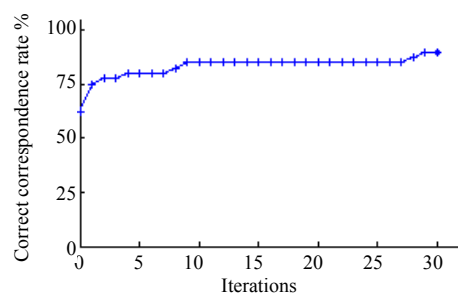
参考文献:

- [1] Kennedy J, Eberhart R C. Particle swarm optimization [C]// Proceedings of IEEE International Conference on Neural Networks. Piscataway, NJ, USA: IEEE Press, 1995, 4: 1942-1948.
- [2] E Ozcan, C K Mohan. Particle swarm optimization: surfing the waves. [C]// IEEE Congress on Evolutionary Computation (CEC 1999), Washington, DC, USA. USA: IEEE, 1999: 1939-1944.
- [3] F van den Bergh. An analysis of particle swarm optimizers [D]. South Africa: Department of Computer Science, University of Pretoria, Nov. 2001.
- [4] M Clerc, J Kennedy. The particle swarm: explosion, stability, and convergence in a multidimensional complex space [J]. IEEE Transactions on Evolutionary Computation (S1089-778X), 2002, 6(1): 58-73.
- [5] Ioan Cristian Trelea. The particle swarm optimization algorithm: convergence analysis and parameter selection [J]. Information Processing Letters (S0020-0190), 2003, 85(66): 317-325.
- [6] M Jiang, Y P Luo, S Y Yang. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm [J]. Information Processing Letters (S0020-0190), 2007, 102(1): 8-16.
- [7] 李宁, 孙德宝, 等. 基于差分方程的 PSO 算法粒子运动轨迹分析 [J]. 计算机学报, 2006, 29(11): 2052-2061.

35 帧图像进行匹配时, LSC 算法的正确率随迭代次数变化的曲线。其中 ‘+’ 表示在迭代过程中仅从行方向进行匹配判定的正确率, ‘◆’ 表示当迭代结束, 将匹配概率矩阵转化为双随机矩阵后, 同时从行和列方向进行判定得到的匹配正确率。从该图可以看出, 匹配的正确率随着迭代过程逐步提高。在迭代的开始阶段, 算法收敛速度很快, 而随着迭代次数的增加, 收敛速度逐渐降低, 这种性质是和概率松弛算法的特点相吻合的。一个很有意义的现象是, 当匹配概率矩阵转化为双随机矩阵后, 匹配正确率并没有降低(第 0 帧和第 30 帧匹配时有所提高), 这进一步说明了用双随机矩阵表示匹配概率矩阵的优越性。LSC 算法对于第 0 帧和第 30 帧, 第 0 帧与第 35 帧进行匹配运算大约需 55 秒。



(a) 第0帧与第30帧匹配



(b) 第0帧与第35帧匹配

图 6 LSC 算法匹配正确率随迭代次数变化曲线

6 结论

本文提出了一种结合形状上下文分析的 Laplace 谱匹配算法。首先, 利用 Laplace 矩阵的特征值与特征向量的组合性质, 并结合双随机矩阵来求解谱匹配概率。其次, 通过一

种松弛迭代的途径, 将形状上下文表示的局部相似性融入 Laplace 谱匹配算法, 以减小随机位置抖动对谱匹配结果的影响。

在本文中, 概率松弛的方法被应用于形状上下文分析和 Laplace 谱匹配算法的结合。需要注意的是, 概率松弛算法只是众多基于约束传播思想的算法中的一种, 如何将神经网络等其他方法应用于形状上下文分析和谱匹配算法的有效结合是我们正在研究的方向之一。另外, 如何从理论上论证本文算法的正确性和完整性也是需要进一步探讨的。

参考文献:

- [1] G L Scott, H C Longuet-Higgins. An Algorithm for Associating the Features of Two Images [C]// Proc. Roy. Soc. London Ser. B (Biological) (S0962-8452). London, UK: Royal Society of London, 1991, 244: 21-26.
- [2] L S Shapiro, J Brady. Feature-based Correspondence: An Eigenvector Approach [J]. Image and Vision Computing (S0262-8856), 1992, 10(5): 283-288.
- [3] M Carcassoni, E R Hancock. Spectral Correspondence for Point Pattern Matching [J]. Pattern Recognit. (S0031-3203), 2003, 36(1): 193-204.
- [4] X Bai, H Yu, E R Hancock. Graph Matching Using Spectral Embedding and Alignment [C]// Proc. 17th Int. Conf. Pattern Recognition (S1063-6919). Los Alamos, USA: IEEE CS Press, 2004, 3: 398-401.
- [5] 王年, 范益政, 韦穗, 梁栋. 基于图的 Laplace 谱的特征匹配[J]. 中国图像图形学报, 2006, 11(3): 332-336.
- [6] R Sinkhorn. A Relationship between Arbitrary Positive Matrices and Doubly Stochastic Matrices [J]. The Annals of Mathematical Statistics (S0003-4851), 1964, 35(2): 876-879.
- [7] S Belongie, J Malik, J Puzicha. Shape Matching and Object Recognition Using Shape Contexts [J]. IEEE Trans. Pattern Anal. and Machine Intell. (S0162-8828), 2002, 24(4): 509-522.
- [8] Y Zheng, D Doermann. Robust Point Matching for Non-rigid Shapes: a Relaxation Labeling Based Approach [J]. IEEE Trans. Pattern Anal. and Machine Intell. (S0162-8828), 2006, 28(4): 643-649.
- [9] H F Wang, E R Hancock. Correspondence Matching Using Kernel Principal Components Analysis and Label Consistency Constraints [J]. Pattern Recognit. (S0031-3203), 2006, 39(6): 1012-1025.
- [10] S Ranade, A Rosenfeld. Point Pattern Matching by Relaxation [J]. Pattern Recognit. (S0031-3203), 1980, 12(4): 269-275.

(上接第 4344 页)

- [8] R C Eberhart, Y Shi. Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization. [C]// The 2000 Congress on Evolutionary Computation (CEC 2000), California, CA, USA, 16 July-19 July 2000: 84-88.
- [9] Y Shi, R C Eberhart. Empirical study of particle swarm optimization. [C]// Proceedings of the 1999 Congress on Evolutionary Computation (CEC 1999) Washington, DC, USA, 07/06/1999-07/09. 1999, 3: 1945-1950.
- [10] Y Shi, R C Eberhart A modified particle swarm optimizer [C]//

Proceedings of IEEE International Conference on Evolutionary Computation. USA: IEEE Service Center, 1998: 69-73.

- [11] Clerc M. The swarm and the queen: towards a deterministic and adaptive particle swarm optimization [C]// Proceedings of the 1999 Congress on Evolutionary Computation (CEC 1999) Washington, DC, USA 07/06/1999 - 07/09. 1999: 1951-1957.
- [12] Zhang L P, Yu H J, Hu S X. Optimal choice of parameters for particle optimization [J]. Journal of Zhejiang University Science (S1673-1581), 2005, 6A(6): 528-534.